



Europäisches
Patentamt

European
Patent Office

10/531013
PCT/IB 03/04497

Office européen
des brevets

13 OCT 2003

REC'D 17 OCT 2003
WIPO PCT

Bescheinigung Certificate

Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

02102461.7

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk



Anmeldung Nr:
Application no.: 02102461.7
Demande no:

Anmeldetag:
Date of filing: 17.10.02
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

Koninklijke Philips Electronics N.V.
Groenewoudseweg 1
5621 BA Eindhoven
PAYS-BAS

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

Einrichtung und Verfahren zum Wiedergeben von Audiodaten sowie
Computerprogrammprodukt hierfür

In Anspruch genommene Priorität(en) / Priority(ies) claimed /Priorité(s)
revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

G10L/

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of
filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LI LU MC NL PT SE SK TR

Einrichtung und Verfahren zum Wiedergeben von Audiodaten
sowie Computerprogrammprodukt hierfür

5 Die Erfindung bezieht sich auf eine Einrichtung zum Wiedergeben von Audiodaten, welche Audiodaten zu Textdaten eines aus Wörtern zusammengesetzten Texts korrespondieren, mit Speichermitteln zum Speichern der Audiodaten, in welche Speichermittel zu speichernde Audiodaten in einer Vorwärtsreihenfolge einspeicherbar sind, und mit Steuermitteln zum Steuern des Wiedergebens von gespeicherten Audiodaten

10 10 in einem Vorwärtsmodus und in einem Rückwärtsmodus, und mit Audio-Wiedergabemitteln.

Die Erfindung bezieht sich weiters auf ein Verfahren zum Wiedergeben von in Speichermitteln gespeicherten Audiodaten, welche Audiodaten zu Textdaten eines aus Wörtern zusammengesetzten Texts korrespondieren und in welche Speichermittel zu speichernde Audiodaten in einer Vorwärtsreihenfolge eingespeichert werden, bei welchem Verfahren das Wiedergeben von Audiodaten in einem Vorwärtsmodus und in einem Rückwärtsmodus steuerbar ist.

Die Erfindung bezieht sich weiters auf ein Computerprogrammprodukt sowie auf einen Computer, der zum Abarbeiten eines solchen Computerprogrammprodukts

20 20 ausgebildet ist.

Beim händischen oder automatischen Transkribieren von Texten, insbesondere beim Korrigieren von automatisch unter Verwendung von Spracherkennungssystemen transkribierten Texten, ist es üblich, den diktieren und in Form von Audiodaten digital gespeicherten Text mit Hilfe von Audio-Wiedergabemitteln, z.B. Kopfhörern, anzuhören, wobei gegebenenfalls, bei bereits transkribierten Texten, die zu korrigieren sind, gleichzeitig der bereits als Textdatei gespeicherte und zu den gespeicherten Audiodaten korrespondierende Text mit Hilfe von Textanzeigemitteln, beispielsweise mit Hilfe von einem Bildschirm eines Computerarbeitsplatzes, angezeigt wird. Insbesondere ist es hierbei auch bekannt, die zueinander korrespondierenden Audiodaten und Textdaten mit entsprechenden Wort-Markierungsdaten zu versehen, die beispielsweise jeweils einen

Wortanfang angeben und die als Verknüpfungsdaten zueinander korrespondierende, also zusammengehörige Audiodaten und Textdaten anzeigen, um diese synchron im Vorwärtsmodus akustisch und optisch wiederzugeben. Eine entsprechende Technik hiefür ist beispielsweise in dem Patentdokument WO 01/46853 A1 beschrieben. Dabei ist es auch

5 bekannt, im optisch anzuzeigenden Textteil jenes Wort, das gerade akustisch wiedergegeben wird, visuell hervorzuheben, was ebenfalls unter Verwendung der durch die Wort-Markierungsdaten bzw. Verknüpfungsdaten gebildeten Steuerdaten realisiert werden kann.

Ein Anhören und Anzeigen von Wörtern in ihrer Relation zueinander ist

10 hierbei nur im Vorwärtsmodus und in der Vorwärtsreihenfolge ermöglicht. Wenn ausgehend von einer momentanen Wiedergabeposition zu einer vorhergehenden Textstelle entgegen der Vorwärtsreihenfolge zurückgesprungen wird, wird gegebenenfalls gleichzeitig auch eine Audio-Wiedergabe ermöglicht, jedoch ebenfalls entgegen der Vorwärtsreihenfolge und somit in einer unverständlichen Form. Wenn beispielsweise ein

15 Diktat automatisch oder aber händisch transkribiert wird und die transkribierende und gegebenenfalls korrigierende Person von der Person, die das Diktat diktiert hat, verschieden ist, wird von dieser das Transkribieren durchführenden Person das Zurückgehen zu vorhergehenden Textstellen als besonders unangenehm empfunden, da ihr der gesprochene Text völlig unbekannt ist, und da sie je nach der verfügbaren Software die

20 Audiodaten - die in digitaler Form in Speichermitteln gespeichert sind - in einer raschen Aufeinanderfolge entgegen der Vorwärtsreihenfolge in unverständlicher Form präsentiert bekommt. Es ist für diese Person dann notwendig, manuell ein Wiedergeben in der Vorwärtsreihenfolge einzuschalten und die betreffenden Diktatteile anzuhören sowie im Fall einer vorhergehenden automatischen Transkription die zugehörigen Textwörter visuell

25 zu überprüfen, wobei nur in diesem Vorwärtsmodus eine zu dem optisch angezeigten Text synchrone Audio-Wiedergabe möglich ist. Dies bedeutet einen vergleichsweise hohen Zeitaufwand, wodurch die Konzentration der die Bearbeitungen durchführenden Person beeinträchtigt werden kann. Außerdem wird auch die Effizienz bei der Bearbeitung nachteilig beeinflusst.

30 In dem Patentdokument US 2002/0062214 A1 ist ein Textmarkierungssystem beschrieben, bei dem Wortgruppen an einem Computer-Bildschirm angezeigt werden, wobei Schaltfelder zur Steuerung verschiedener Arbeitsschritte vorgesehen sind. Dabei

sind zwei gesondert anzusteuernde Schaltfelder vorgesehen, um ausgehend von einem markierten, optisch in einer Textzeile hervorgehobenen Wort zum unmittelbar davor stehenden Wort oder zum unmittelbar danach stehenden Wort zu springen, um dieses Wort optisch hervorzuheben und gleichzeitig akustisch wiederzugeben. Diese Steuerung ist

5 jedoch dann, wenn ausgehend von einem bestimmten Wort eine Textstelle relativ weit davor, zum Beispiel 10 oder 20 Wörter davor, gesucht wird, außerordentlich mühsam und zeitaufwendig, wobei immer wieder ein händisches "Anklicken" des entsprechenden Schaltfeldes notwendig ist.

10

Die Erfindung hat sich zur Aufgabe gestellt, hier Abhilfe zu schaffen und eine Einrichtung und ein Verfahren zu realisieren, die ein rasches und zielgerechtes Aufsuchen von gesprochenen Textteilen in gespeicherten Audiodaten ermöglichen, wobei auch möglichst wenige händische Steuereingriffe seitens der die Bearbeitung durchführenden

15 Person erforderlich sein sollen.

Gemäß einem ersten Aspekt sieht die Erfindung zur Lösung der vorstehend angeführten Aufgabe eine Einrichtung zum Wiedergeben von gespeicherten Audiodaten vor, welche Audiodaten zu Textdaten eines aus Wörtern zusammengesetzten Texts korrespondieren, mit Speichermitteln zum Speichern der Audiodaten, in welche

20 Speichermittel zu speichernde Audiodaten in einer Vorwärtsreihenfolge einspeicherbar sind, und mit Steuermitteln zum Steuern des Wiedergebens von gespeicherten Audiodaten in einem Vorwärtsmodus und in einem Rückwärtsmodus, und mit Audio-Wiedergabemitteln, wobei die Steuermittel derart eingerichtet sind, dass sie bei einem Wiedergeben von Audiodaten im Rückwärtsmodus ausgehend von der jeweiligen

25 momentanen Wiedergabeposition in den Audiodaten entgegen der Vorwärtsreihenfolge selbsttätig ein Zurückspringen um eine Rücksprunglänge, die der Länge von mindestens etwa zwei Wörtern entspricht, zu einer Zielposition veranlassen und danach beginnend bei der jeweiligen Zielposition ein Wiedergeben von Audiodaten in der Vorwärtsreihenfolge für nur einen Teil der Rücksprunglänge veranlassen.

30 Gemäß einem zweiten Aspekt sieht die Erfindung ein Verfahren zum Wiedergeben von in Speichermitteln gespeicherten Audiodaten vor, welche Audiodaten zu Textdaten eines aus Wörtern zusammengesetzten Text korrespondieren und in welche

Speichermittel zu speichernde Audiodaten in einer Vorwärtsreihenfolge eingespeichert werden, bei welchem Verfahren das Wiedergeben von Audiodaten in einem Vorwärtsmodus und in einem Rückwärtsmodus steuerbar ist, wobei bei einem Wiedergeben von Audiodaten im Rückwärtsmodus ausgehend von der jeweiligen

5 momentanen Wiedergabeposition in den Audiodaten selbsttätig um eine Rücksprunglänge, die der Länge von mindestens etwa zwei Wörtern entspricht, entgegen der Vorwärtsreihenfolge zu einer Zielposition zurückgesprungen wird und danach beginnend bei der jeweiligen Zielposition ein Wiedergeben in der Vorwärtsreihenfolge für nur einen Teil der Rücksprunglänge durchgeführt wird.

10 Mit dem erfindungsgemäßen Verfahren kann das Suchen bestimmter Textpassagen in den Audiodaten im Vergleich zum Stand der Technik rascher und effizienter durchgeführt werden. Wenn sich beispielsweise bei einem durch eine Person durchgeführten Transkribieren oder Korrigieren eines Textes beim Erreichen einer Textstelle das Problem ergibt, dass diese Person bei einer zuvor transkribierten oder

15 korrigierten Textstelle, die um 10 oder 20 Wörter im zu transkribierenden oder zu korrigierenden Text vorher aufgetreten ist, möglicherweise eine Unklarheit oder eine Unstimmigkeit oder einen Fehler erkennt, so kann außerordentlich rasch und völlig automatisch - nach Starten des Verfahrens gemäß der Erfindung - ein entsprechendes Suchen im Rückwärtsmodus durchgeführt werden, wobei rechnergestützt automatisch

20 entsprechend den vorgegebenen Rücksprunglängen weiter zurückliegende Zielpositionen im Text angesprungen werden und danach für nur einen vorgegebenen Teil der jeweiligen Rücksprunglänge in der Vorwärtsreihenfolge eine akustische Wiedergabe durchgeführt wird. Dadurch wird eine verständliche Audio-Wiedergabe realisiert, so dass die jeweilige Person keine Verständnisschwierigkeiten hat. Das Zurückspringen im durch die

25 Audiodaten gegebenen Text erfolgt entweder, wenn schon ein korrespondierender transkribierter Text vorliegt, bevorzugt unter Verwendung von Wort-Markierungsdaten als Steuerzeichen, welche üblicherweise jeweils den Beginn eines neuen Wortes angeben. Falls noch kein transkribierter Text vorliegt, wird die Rücksprunglänge für das automatische Zurückspringen in den Audiodaten z.B. in Entsprechung zu der mittleren

30 Datenlänge oder Zeit von zumindest zwei Wörtern geschätzt, wobei beispielsweise eine Vorwärts-Wiedergabezeit von ein oder zwei Sekunden zugrundegelegt wird. In diesem Zusammenhang sei daher auch festgehalten, dass die jeweilige Rücksprunglänge nicht

exakt der Länge von mehreren tatsächlich gesprochenen Wörtern entsprechen muss, da die Audiodaten auch in Segmente entsprechend gemittelten "Wort-Längen" unterteilt werden können. Ähnliches gilt für die Dauer des nach jedem Zurückspringen vorgesehenen Wiedergebens in der Vorwärtsreihenfolge.

5 Es erfolgt demgemäß bei dem erfindungsgemäßen Verfahren beispielsweise jeweils ein Sprung um (ungefähr) zwei oder drei Wörter zurück und danach eine automatische Audio-Wiedergabe jeweils (ungefähr) eines Wortes, wobei das gerade wiedergegebene Wort danach eines der zwei oder drei Wörter ist, um die im jeweils nachfolgenden Verfahrensschritt in den Audiodaten zurückgesprungen wird. Denkbar ist es

10 jedoch auch, jeweils um eine Rücksprunglänge entsprechend einer größeren Anzahl von Wörtern zurückzuspringen und beim Wiedergeben in der Vorwärtsreihenfolge immer nur ein Wort im gesprochenen Text anzusteuern bzw. wiederzugeben, so dass beispielsweise nur jedes vierte oder fünfte Wort wiedergegeben wird. Es können aber auch mehrere Wörter in der Vorwärtsreihenfolge wiedergegeben werden.

15 Die Audio-Wiedergabe kann jeweils einfach wortweise erfolgen, das heißt bis zum Erscheinen der nächsten Wort-Markierungsdaten, wenn diese bereits verfügbar sind, wobei diese nächsten Wort-Markierungsdaten den Beginn des nächsten Wortes kennzeichnen. Es kann jedoch auch eine segmentweise Wiedergabe mit einer festen Wiedergabezeit, beispielsweise im Bereich von 0,5 bis 1,5 Sekunden, z.B. entsprechend einer gemittelten Wortedauer, durchgeführt werden, die von einem Zeitsteuerkreis vorgegeben wird. Dies wird vor allem dann der Fall sein, wenn noch kein transkribierter Text, mit entsprechenden Wort-Markierungsdaten, vorliegt. Auch ist es denkbar, jeweils ein Wort zur Gänze und das nächste markierte Wort nur teilweise am Wortanfang wiederzugeben. Ferner ist es denkbar, dass die die Bearbeitung durchführende Person 20 zwischen den verschiedenen vorstehend angeführten Möglichkeiten wählt.

25 Weiters wird erfindungsgemäß auch bevorzugt, dass beim beschriebenen Rückwärts-Suchlauf die jeweilige Audio-Wiedergabe in der Vorwärtsreihenfolge mit einer einstellbaren Geschwindigkeit erfolgt, um seitens der die Bearbeitung durchführenden Person den Suchlauf insgesamt je nach Notwendigkeit rascher oder langsamer ablaufen zu lassen. Das Zurückspringen zur jeweiligen früheren Zielposition im wiederzugebenden Text, wie sie z.B. durch die erwähnten Wort-Markierungsdaten vorgegeben sein kann, kann ganz rasch, also praktisch ohne Zeitverlust, durchgeführt werden, wobei bei diesem

"raschen Rücklauf"-Modus auch keine akustische Audiowiedergabe durchgeführt zu werden braucht.

Von ganz besonderem Vorteil ist das erfindungsgemäße Verfahren in Verbindung mit einem Transkriptionssystem einsetzbar, bei dem in an sich herkömmlicher

5 Weise beispielsweise über ein Kommunikationsnetz, z.B. LAN, WAN oder Internet, oder über Tonträger einlangende Diktate automatisch mit Hilfe von Spracherkennungsmitteln in eine Textdatei umgesetzt werden, die sodann mit Hilfe einer Textverarbeitungssoftware unter Anhören der Diktat-Audiodaten überprüft und gegebenenfalls korrigiert wird. Dabei ist eine Verknüpfung der Wörter in der Audiodatei einerseits und der zugehörigen Wörter

10 10 in der Textdatei andererseits aufgrund der zugeordneten Wort-Markierungsdaten gegeben, die deshalb auch als Verknüpfungsdaten bezeichnet werden. Bei der Wiedergabe wird auf den Textanzeigemitteln das jeweils gerade akustisch wiedergegebene Wort auch optisch hervorgehoben, beispielsweise durch Unterlegen mit einem hellen Hintergrund (sog. "Highlighten"). Hier sieht die Erfindung einen "synchronen Rückwärts-Wiedergabe-

15 15 Modus" vor, wobei die Wörter der Textdatei der Reihe nach - entgegen der Vorwärtsreihenfolge - optisch hervorgehoben werden und synchron zu jedem optisch hervorgehobenen Wort das zu diesem Wort korrespondierende Wort in den Audiodaten in der Aufzeichnungsreihenfolge, also verständlich, akustisch wiedergegeben wird. Dadurch wird der Vorteil erzielt, dass ein Prüfen von zusammenhängenden, optisch angezeigten

20 20 Wörtern anhand einer verständlichen Audio-Wiedergabe der korrespondierenden Audiodaten mühelos durchführbar ist. Auch wird ein Lokalisieren einer Position im Text wesentlich vereinfacht, und insgesamt wird die Effizienz beim Korrigieren von transkribierten Diktaten erhöht.

Die Erfindung kann somit in einem solchen typischen Transkriptionssystem

25 25 mit Vorteil angewendet werden, bei dem Diktate mit Hilfe von Empfangsstationen empfangen und mit Hilfe von Transkriptionsstationen automatisch transkribiert werden, wonach mit Hilfe von Korrekturstationen ein manuelles Korrigieren der transkribierten Diktate und schließlich mit Hilfe von Abgabestationen ein Abgeben von Textdateien entsprechend den empfangenen Diktaten durchgeführt wird. Selbstverständlich kann aber

30 30 die Erfindung auch bei einem Transkriptionssystem eingesetzt werden, das durch einen einzigen Rechner, insbesondere Personalcomputer, realisiert ist, mit dem die erwähnten Schritte, nämlich das Empfangen, das automatische Transkribieren, das Korrigieren und

schließlich das Ausgeben der Textdaten durchgeführt werden können.

Wie bereits angedeutet kann die Erfindung überdies auch bei einem manuellen Transkribieren eines diktierteren Textes eingesetzt werden, wenn anhand des Diktates, also der Audiodaten, die angehört werden, manuell an einem Textverarbeitungssystem die Textdatei, vorzugsweise unter Zuordnen von Verknüpfungsdaten für die Audiodaten und die Textdaten entsprechend den vorstehend erwähnten Wort-Markierungsdaten, erstellt wird, wobei nach dem Erstellungs- oder Transkriptionsvorgang gegebenenfalls auch Korrekturen durchgeführt werden. Insbesondere kann daher die Erfindung auch in einem mobilen Diktier- bzw. Audio-Wiedergabegerät in digitaler Bauweise realisiert werden.

10 Gemäß einem dritten Aspekt sieht die Erfindung auch ein Computerprogrammprodukt vor, das in einen Speicher eines Computers geladen werden kann und Softwarecodeabschnitte umfasst, um durch deren Abarbeiten nach dem Laden in den Speicher des Computers mit dem Computer das erfindungsgemäße Verfahren durchführen zu können.

15 Gemäß einem vierten Aspekt sieht die Erfindung schließlich einen Computer mit einer Recheneinheit und einem internen Speicher vor, welcher Computer zum Abarbeiten des erfindungsgemäßen Computerprogrammprodukts ausgebildet ist.

20 Die Erfindung wird nachfolgend anhand von bevorzugten Ausführungsbeispielen, auf die sie jedoch nicht beschränkt sein soll, und unter Bezugnahme auf die in den beiliegenden Zeichnungen enthaltenen Figuren noch weiter erläutert.

25 Die Figur 1 zeigt schematisch einen Ablauf bei einem synchronen Wiedergeben von Audiodaten und Textdaten in einem Vorwärtsmodus.

Die Figur 2 zeigt schematisch einen Ablauf bei einem Wiedergeben von Audiodaten und Textdaten unter gegenseitiger Zuordnung in einem Rückwärtsmodus gemäß dem Stand der Technik.

30 Die Figur 3 zeigt schematisch einen Ablauf bei einem Wiedergeben von Audiodaten und Textdaten in einem Rückwärtsmodus gemäß der Erfindung.

Die Figur 4A zeigt einen dem Ablauf gemäß der Figur 3 ähnlichen Ablauf zum hörbaren Wiedergeben von Audiodaten in einem Rückwärtsmodus, wobei ein zuvor

automatisch transkribierter und zu korrigierender Text veranschaulicht ist.

Die Figur 4B zeigt den anhand des in der Figur 4A gezeigten Ablaufs bei einem Wiedergeben im Rückwärtsmodus korrigierten Text als beispielsweise auf einem Bildschirm angezeigte Abfolge von Wörtern.

5 Die Figur 5 zeigt schematisch in Form eines Blockschaltbilds ein Transkriptionssystem mit einer Einrichtung zur Audio-Wiedergabe, mit der "synchrone Rückwärtsmodus-Wiedergabe" gemäß den schematischen Darstellungen gemäß den Figuren 3 und 4A durchführbar ist.

Die Figur 6 zeigt in Form von einem Blockschaltbild und auf detailliertere 10 Weise die für eine „synchrone Rückwärtsmodus-Wiedergabe“ vorgesehenen Systemteile des Transkriptionssystems gemäß der Figur 5.

Die Figur 7 zeigt einen dem Ablauf gemäß der Figur 4A ähnlichen, jedoch etwas vereinfacht dargestellten modifizierten Ablauf bei einem Wiedergeben im Rückwärtsmodus.

15 Die Figur 8 zeigt in Form von einem Ablaufdiagramm eine Variante für ein Verfahren zum synchronen Wiedergeben im Rückwärtsmodus.

In der Figur 1 ist schematisch ein Ablauf beim Wiedergeben von Audiodaten 20 A1..A4 (allgemein Ai) synchron mit Textdaten T1..T4 (allgemein Ti) in einem Vorwärtsmodus veranschaulicht, wobei das Wiedergeben (Auslesen) der Daten in der gleichen Reihenfolge bzw. in der gleichen Richtung bzw. in dem gleichen Sinn wie beim Aufzeichnen (Einspeichern) der Daten (in Figur 1 von links nach rechts) erfolgt. Diese Reihenfolge wird stets als Vorwärtsreihenfolge bezeichnet. Die Audiodaten Ai und 25 Textdaten Ti repräsentieren dabei in zusammengehörenden Paaren eine Aufeinanderfolge von Wörtern A1/T1..A4/T4 eines Textes. Dem jeweiligen Wortanfang ist ein Wort-Markierungscode oder sind Wort-Markierungsdaten M1..M5 (allgemein Mi) zugeordnet, die zugleich Verknüpfungsdaten für eine synchrone Wiedergabe der Audiodaten Ai und Textdaten Ti bilden. Bei der Wiedergabe werden die Textdaten T1, T2...T4 (also 30 aufeinander folgende Wörter) entsprechend den in Figur 1 unten gezeigten Pfeilen 1, 2, 3, 4 aufeinander folgend angesteuert und optisch auf einem Anzeigemittel (in Figur 1 nicht dargestellt) hervorgehoben, und synchron dazu wird aus den entsprechenden digitalen

Audiobanden A1, A2...A4 das jeweilige Wort entsprechend den oberhalb hiervon mit Pfeilen 1, 2, 3 und 4 angegebenen Schritten akustisch wiedergegeben. Dieses gleichzeitige optische und akustische Wiedergeben von Wörtern des Textes unter Nutzung von Markierungs- bzw. Verknüpfungsdaten Mi in einem Vorwärtsmodus ist an sich bekannter Stand der

5 Technik.

In der Figur 2 ist schematisch ein ebenso bekannter Ablauf beim Wiedergeben in einem Rückwärtsmodus veranschaulicht. Hier werden die Wörter T4, T3, T2, T1 von hinten nach vorne, gemäß der Darstellung in Figur 2 von rechts nach links, aufeinander folgend gemäß den unteren Pfeilen 1, 2, 3 und 4 angesteuert und auf dem nicht gezeigten

10 Anzeigemittel visuell hervorgehoben. Gleichzeitig werden unter Verwendung der Wort- Markierungs- oder Verknüpfungsdaten M5, M4, M3 und M2 die entsprechenden

15 Audiobanden A4, A3, A2 und A1, das heißt die Wörter entgegen der Vorwärtsreihenfolge gemäß den oberen Pfeilen 1, 2, 3 und 4 in Figur 2, wiedergegeben. Dieses akustische Wiedergeben erfolgt somit entgegen der Aufzeichnungsreihenfolge, also entgegen der

15 Aufzeichnungsrichtung, und hat daher ein unverständliches Audiosignal zur Folge.

Dadurch ist das Auffinden bestimmter Textstellen erschwert und nur aufgrund der visuellen Anzeige möglich, was jedoch der üblichen Arbeitsweise beim Transkribieren bzw. Korrigieren von Texten nach einem Diktat entgegensteht, da Personen, die eine Transkription oder die Korrektur einer Transkription durchführen, beim Rückwärtssuchen

20 ebenso wie beim Vorwärtssuchen ihre Konzentration auf ein akustisch wiedergegebenes Audiosignal richten, wobei diese Personen ja auch beim direkten Bearbeiten entsprechend dem gehörten Audiosignal den Text schreiben bzw. korrigieren.

25 Sofern noch kein transkribierter Text (mit Wörtern T1..T4..) vorliegt, ist in diesem Fall das Auffinden von weiter zurückliegenden Wörtern unter alleiniger Ausnutzung der Audiobanden Ai überhaupt äußerst mühsam.

30 Anders als bei diesen bekannten Techniken wird nun vorgesehen, auch im Rückwärtsmodus die einzelnen Audiobanden Ai, beispielsweise in der vorstehenden Reihenfolge A4, A3, A2 und A1, jeweils in der Vorwärtsreihenfolge, das heißt in der Aufzeichnungsrichtung bzw. in dem Aufzeichnungssinn, akustisch wiederzugeben, wie dies schematisch in Figur 3 mit den Pfeilen 1, 2, 3, 4 oberhalb der Audiobanden A4, A3, A2, A1 veranschaulicht ist. Gleichzeitig wird, wenn entsprechende Textdaten T4, T3, T2, T1 bereits vorhanden sind, eine visuelle Anzeige der durch die Textdaten Ti repräsentierten

Wörter gemäß den Pfeilen 1, 2, 3 und 4 in Figur 3 unten veranlasst.

In der Figur 4A ist nunmehr im Detail veranschaulicht, wie bei einem derartigen "synchrone[n] Wiedergeben im Rückwärtsmodus" in den Audiodaten A1 bis A6 ebenso wie in den zugehörigen Textdaten T1 bis T6 zurückgesprungen wird und jeweils 5 eine verständliche Audio-Wiedergabe erzeugt wird. Es wird dabei als Beispiel von einer Textpassage in einem Diktat ausgegangen, die richtig "TO BE OR NOT TO BE" (siehe auch Figur 4B) lautet, die jedoch von einem automatischen Transkriptionssystem in der Form wie in der Figur 4A dargestellt transkribiert wurde, nämlich als "TWO BEE OR NOT TWO BEE". In der Figur 4A ist diese Wortfolge in einem Balken 11 veranschaulicht, 10 der als solcher beispielsweise auf einem nicht gezeigten optischen Anzeigemittel, z.B. einem Bildschirm, erscheint, um so die einzelnen Wörter - die Textdaten T1, T2..T6 - visuell anzuzeigen. Diese Wörter sind auch als entsprechende Audiodaten A1, A2..A6 in digitaler Form in der Figur 4A nicht näher gezeigten Audiodaten-Speichermedien gespeichert und aus diesen zur Audio-Wiedergabe auslesbar. Die hiefür vorgesehenen 15 Wort-Markierungsdaten bzw. Verknüpfungsdaten Mi sind in der Figur 4A wiederum schematisch bei M1, M2.....M7 angedeutet.

Konkret wird nun gemäß der Figur 4A von einer im gesprochenen bzw. transkribierten Text Ti weiter hinten (in Figur 4A weiter rechts) liegenden momentanen Wiedergabeposition zu einer vorhergehenden Zielposition zurückgesprungen, z.B. zum 20 durch die Verknüpfungsdaten M6 gekennzeichneten Anfang des Wortes T6 bzw. A6 ("BEE"). Dieses Zurückspringen ist in der Figur 4A mit einem Pfeil 1A angedeutet. Danach wird beispielsweise genau dieses Wort A6 aus den gespeicherten Audiodaten Ai in der Vorwärtsreihenfolge wiedergegeben, vergleiche den Pfeil 1B. Am Ende des Wortes A6/T6 (bzw. bei einem durch Markierungsdaten M7 markierten nächsten Wort) wird dann 25 selbsttätig, und zwar in diesem Fall mit einer Mindest-Rücksprunglänge, die der Länge von zwei Wörtern A5 + A6 bzw. T5 + T6 entspricht, zum durch die Verknüpfungsdaten M5 gekennzeichneten Anfang des Wortes T5 (Textdaten) bzw. A5 (Audiodaten) zurückgesprungen (siehe Pfeil 2A), wonach das Wort A5 in der Vorwärtsreihenfolge gemäß dem Pfeil 2B als Audiosignal wiedergegeben wird. Dieser Vorgang wird sukzessive 30 automatisch mit den Wörtern A4/T4, A3/T3 usw. wiederholt, vergleiche die Pfeile 3A (Zurückspringen zur Zielposition M4), 3B (akustische Wiedergabe von dem Wort A4 in Aufzeichnungsrichtung) usw. bis zu den Pfeilen 6A, 6B. Die Pfeile 1A, 2A, 3A...6A geben

in der Figur 4A demgemäß die Rücksprunglängen an, wogegen die Pfeile 1B, 2B, 3B...6B allgemein diejenigen Teile der Rücksprunglängen angeben, für die eine Audio-Wiedergabe in der Vorwärtsreihenfolge erfolgt.

Während des beschriebenen abschnittsweisen Zurückspringens und Anhörens

- 5 der einzelnen Wörter Ai in der Vorwärtsreihenfolge kann an sich das jeweilige Wort unmittelbar korrigiert werden, oder es wird bis zum Beginn der betroffenen Textpassage zurückgekehrt, wonach in üblicher Weise die gesamte Wortfolge T1 bis T6 durch Anhören in der Vorwärtsreihenfolge und optisches Anzeigen ebenfalls in der Vorwärtsreihenfolge korrigiert werden kann, so dass der richtige Text, gemäß dem Balken 11' in der Figur 4B,
- 10 erhalten wird.

Während der jeweiligen akustischen Wiedergabe der einzelnen Audiodaten A1..A6 werden die dazu korrespondierenden Textdaten T1..T6 am Bildschirm optisch besonders hervorgehoben, beispielsweise durch Anzeigen eines hellen Hintergrundes (sogenanntes "Highlighten").

- 15 In der Figur 4A ist schematisch beim Pfeil 2B noch durch eine strichlierte Verlängerung angedeutet, dass die akustische Wiedergabe jeweils über das bestimmte Wort hinausgehend auch das nächstfolgende Wort - teilweise - erfassen kann, das heißt, dass eine "wortüberlappende" Audio-Wiedergabe vorgesehen sein kann. Dies ergibt sich beispielsweise dann, wenn zum Beenden des jeweiligen Audio-Wiedergabeschritts nicht
- 20 die jeweils nächstfolgenden Markierungsdaten Mi, zum Beispiel M6, als Steuercode verwendet werden, sondern wenn eine feste Wiedergabezeit aufgrund einer Zeitzählung vorgesehen ist. Die feste Wiedergabezeit kann beispielsweise eine Sekunde oder 1,5 Sekunden, eventuell auch etwas unter einer Sekunde betragen. Eine solche vorgegebene Zeitdauer für die Audio-Wiedergabe ist insbesondere dann vorzusehen, wenn noch kein
- 25 transkribierter Text vorliegt und somit auch noch keine Wort-Markierungsdaten als Steuercode verfügbar sind.

In ähnlicher Weise kann dann auch die Rücksprunglänge beim Zurückspringen in Entsprechung zu fix vorgegebenen Zeitspannen, z.B. entsprechend dem Zwei- oder Dreifachen einer mittleren Wortlänge, bemessen werden.

- 30 In der Figur 5 ist ein Beispiel für eine Einrichtung 12 zum Wiedergeben von Audiodaten Ai synchron mit dem Wiedergeben von Textdaten Ti dargestellt, die eine Transkriptions-Datenverarbeitungseinrichtung 13 enthält. In an sich üblicher Weise wird

bei dieser Einrichtung 12 von Benutzern 14.1.....14.N über ein Kommunikationsmedium, zum Beispiel über ein Kommunikationsnetz 15, wie LAN, WAN oder Internet, eine Diktatdatei an die Einrichtung 12 übermittelt, die über eine Kommunikationseinrichtung, im vorliegenden Fall ein Modem 16, empfangen und dann Spracherkennungsmitteln 17 5 zugeführt wird. Es sei erwähnt, dass die Kommunikationsmittel auch durch ein sogenanntes "Private Branch Exchange", kurz PBX, realisiert sein können.

Die Spracherkennungsmittel 17, in denen eine an sich übliche Spracherkennungssoftware abgearbeitet wird, führen eine automatische Transkription der Diktatdatei in eine Textdatei durch, wobei die einzelnen Wörter der Textdaten Ti unter 10 Generierung von Wort-Markierungsdaten bzw. Verknüpfungsdaten Mi in Entsprechung zu den einzelnen Audiodaten Ai der zugehörigen Audiodatei abgespeichert werden, und zwar in Text-Speichermitteln 18 bzw. in Audio-Speichermitteln 19. Durch die Wort-Markierungsdaten Mi werden, wie vorstehend bereits ausgeführt, die jeweils einander entsprechenden bzw. jeweils zueinander korrespondierenden Wörter in den Textdaten Ti 15 bzw. in den Audiodaten Ai (Text-Speichermittel 18 bzw. Audio-Speichermittel 19) einander fest zugeordnet, das heißt miteinander verknüpft. Über diese Verknüpfung können durch Steuermittel 20 zueinander gehörende Audiodaten und Textdaten Ai und Ti paarweise aufgerufen und wiedergegeben werden. Die von den Steuermitteln 20 veranlasste visuelle Wiedergabe der Textdaten Ti erfolgt über Textverarbeitungsmittel 21 20 an Anzeigemitteln 22, wie insbesondere an einem Computer-Bildschirm.

Die akustische Wiedergabe der Audiodaten Ai erfolgt unter Auslesen der digital gespeicherten Audiodaten Ai aus den Speichermitteln 19 und Zuführung derselben zu einem Wiedergabekreis 23 für einen elektroakustischen Wandler 24, wobei in der Regel ein Kopfhörer hiefür verwendet wird. Das Auslesen erfolgt hierbei in der 25 Vorwärtsreihenfolge.

Die Steuerung des gesamten Ablaufs beim Zurückspringen von der jeweiligen momentanen Wiedergabeposition zu vorhergehenden Zielpositionen im Text und zur synchronen Vorwärts-Wiedergabe erfolgt anhand von in einem internen Speicher 25 gespeicherten Softwarecodeabschnitten. Als Schnittstelle zum Benutzer, und zwar zur 30 Aktivierung der jeweiligen Steuervorgänge sowie für die diversen Eingaben im Zuge der Transkription bzw. Textkorrektur, dient eine herkömmliche Tastatur oder dergleichen als Eingabemittel 26. Es kann auch eine Fußschalter-Bedienungseinrichtung zum Steuern der

Vorwärts- und Rückwärts-Wiedergabe vorgesehen sein.

Der Wiedergabeschaltkreis 23, der in an sich üblicher Weise einen Digital/Analog-Konverter, einen Verstärker und dergl. Komponenten enthalten kann, und der Wandler 24 bilden zusammen Audio-Wiedergabemittel 27.

5 In der Figur 6 ist mehr im Detail dargestellt, wie die Steuermittel 20 das Wiedergeben bzw. das Auslesen von Audiodaten A_i aus den Audio-Speichermitteln 19 und das Zuführen zu den Wiedergabemitteln 27 in Zuordnung zu den Textdaten T_i samt Wort-Markierungsdaten M_i , die z.B. in den Text-Speichermitteln 18 gespeichert sind, steuern. Dabei ist ein zentraler Steuerkreis 28 direkt bzw. über Festlegungsmittel 29, die die 10 jeweilige Rücksprunglänge beim Zurückspringen in den Audiodaten A_i festlegen, mit den Audio-Speichermitteln 19 verbunden. Weiters ist mit dem Steuerkreis 28 ein Zeitsteuerkreis 30 verbunden, um beim Rückwärtsmodus wie vorstehend anhand der Figur 4A beschrieben ein akustisches Wiedergeben der Audiodaten in der Vorwärtsreihenfolge über eine vorgegebene, feste Zeitdauer zu ermöglichen. Der Zeitsteuerkreis 30 kann zum 15 Beispiel ein üblicher Taktgenerator sein, wobei der Steuerkreis 28 die jeweils gewünschte Zeitdauer - die beispielsweise über die Eingabemittel 26 eingestellt werden kann - durch entsprechendes Abzählen von Taktimpulsen misst. Auf diese Weise kann eine Wiedergabe-Zeitdauer (vergleiche Pfeil 2B in Figur 4A) von beispielsweise einer Sekunde oder 1,5 Sekunden eingestellt werden. Als Alternative kann, wenn bereits Wort- 20 Markierungsdaten M_i vorhanden sind, auch eine "wortweise" Wiedergabe gewählt werden, wobei bei Erreichen der jeweils nächstfolgenden Markierungsdaten M_i der Wiedergabevorgang durch den Steuerkreis 28 beendet wird. Vom Zeitsteuerkreis 30 abgegebene Zeitsteuersignale, insbesondere Taktimpulse, können auch als Basis für die 25 Bestimmung der Rücksprunglängen in den Festlegungsmitteln 29 verwendet werden. Weiters ist es bevorzugt möglich, die Rücksprunglänge beim abschnittsweisen Zurückspringen zu Zielpositionen im vorhergehenden Text einzustellen, wobei ein Überspringen von jeweils mehr als zwei Wörtern, zum Beispiel von drei, vier oder fünf Wörtern, denkbar ist, und wobei entsprechende Anzahl von Markierungsdaten oder Verknüpfungsdaten M_i abzuzählen sind. Zu diesem Zweck ist der Steuerkreis 28 mit 30 Zählmitteln 31 verbunden.

Selbstverständlich können diese Zählmittel 31 ebenso wie der Zeitsteuerkreis 30 softwaremäßig mit Hilfe der Steuermittel 20 realisiert sein, ebenso wie die

Festlegungsmittel 29, die durch eine entsprechende Adressierung in den Speichermitteln 19 realisiert werden können. Es fällt dann der Steuerkreis 28 gemäß der Figur 6 mit den Steuermitteln 20 gemäß der Figur 5 zusammen. Zusätzlich sind noch Einstellmittel 32 zur Einstellung der Geschwindigkeit der Audio-Wiedergabe mit den Wiedergabemitteln 27 vorgesehen.

Es sei erwähnt, dass die Wort-Markierungsdaten Mi zur Gänze auch in den Audio-Speichermitteln 19 bzw. aufgeteilt sowohl in den Text-Speichermitteln 18 als auch in den Audio-Speichermitteln 19 gespeichert sein können.

In der Figur 7 ist zwecks besserer Veranschaulichung schematisch die

10 Vorgangsweise beim Zurückspringen um jeweils mehr als zwei Wörter, beispielsweise um jeweils drei Wörter, veranschaulicht, vergleiche den Pfeil 1A in Rückwärtsrichtung, wobei nach einem solchen Rückwärts-Sprungvorgang eine Audio-Wiedergabe des jeweils folgenden Worts in Vorwärtsrichtung, also der Vorwärtsreihenfolge, erfolgt, siehe beispielsweise Pfeil 1B in der Figur 7. Auf diese Weise werden im Beispiel gemäß der

15 Figur 7 nur jene Wörter akustisch wiedergegeben, die mit den Markierungsdaten M7, M5, M3, M1 (in dieser Reihenfolge) gekennzeichnet sind, vergl. auch den Punkt über diesem Bezugszeichen in der Figur 7. Jene Wörter, denen die Markierungsdaten M8, M6, M4 und M2 zugeordnet sind, werden jedoch hinsichtlich einer akustischen Wiedergabe übersprungen.

20 Selbstverständlich können auch mehr als jeweils drei Wörter beim Zurückspringen übersprungen werden, so dass dann bei der jeweils darauffolgenden akustischen Wiedergabe nur jedes dritte, vierte usw. Wort wiedergegeben wird.

In der Figur 8 ist ein Ablaufdiagramm dargestellt, anhand von dem beispielhaft die Vorgangsweise bei der vorstehend erläuterten synchronen Rückwärtswiedergabe und

25 Vorwärtswiedergabe erläutert werden soll. Das Ablaufdiagramm umfasst dabei auch Konfigurationsschritte und Berechnungsschritte, die dem eigentlichen Rückwärtssprung-Vorgang und Vorwärtswiedergabe-Vorgang vorhergehen.

Gemäß der Figur 8 erfolgt einleitend bei einem Block 33 eine Konfiguration von Optionen O1 bis O5 hinsichtlich der Audiowiedergabe und der Bemessung der

30 Rücksprunglängen 1A, 2A, 3A...6A bei Zurückspringen in den Audiodaten, wie vorstehend beschrieben wurde. Diese Optionen O1 bis O5 können beispielsweise wie folgt vorgesehen sein:

O1 - Diese Option O1 wird dann gewählt, wenn Spracherkennungsmittel 17 vorgesehen und benutzt werden, um automatisch Textdaten T_i aus den Audiodaten A_i zu erstellen, wobei dann auch die beschriebenen Wort-Markierungsdaten M_i von den Spracherkennungsmitteln 17 als Verknüpfungsdaten selbsttätig festgelegt werden.

5 O2 - Diese Option O2 betrifft den Fall, dass manuell transkribierte Textdaten T_i verwendet werden sollen, um die Länge einzelner Segmente oder "Wörter" der Audiodaten A_i festzulegen. Dabei wird gemäß dieser Option O2 eine feste Länge für alle Audio-Segmente vorgesehen, welche Länge mit Hilfe der Steuermittel 20 aus der Gesamt-Zeitdauer der Audiodaten A_i und der Anzahl der transkribierten Wörter durch einfaches

10 Dividieren berechnet wird. Die einzelnen Audio-Segmente oder -Wörter können dann "nummeriert", d.h. mit Adressen bzw. Indizes versehen werden, um anhand der Nummern oder Adressen die Rücksprunglängen 1A, 2A, 3A...6A bzw. die akustisch wiederzugebenden Teile 1B, 2B, 3B...6B zu bestimmen.

15 O3 - Diese der Option O2 sehr ähnliche Option O3 wird gewählt, wenn - ebenfalls auf Basis eines händischen Transkribierens der Textdaten T_i aufgrund der gehörten Audiodaten A_i - von den Steuermitteln 20 die Audio-Segmente mit variabler Länge berechnet werden, und zwar auf Basis der Silben der Wörter und der Gesamtlänge der Audiodaten A_i . Dabei werden alle Silben als gleich lang angenommen.

20 O4 - Die Option O4 kann dann gewählt werden, wenn noch keine Textdaten T_i vorliegen, wobei die Länge von Audio-Segmenten oder Wörtern der Audiodaten A_i auf der Basis von Audioenergieprofil-Informationen berechnet werden.

25 O5 - Bei dieser Option O5 werden Audio-Segmente mit festen Längen angenommen, wie z.B. mit Segmentlängen von einer Sekunde, wobei auch beispielsweise eine Überlappung der Segmente, etwa mit einer Zeitdauer von 1/3 Sekunde, vorgesehen werden kann.

Weiters können beim Block 33 auch Geschwindigkeiten für die normale Wiedergabe in der Vorwärtsreihenfolge, für das rasche Zurückspringen und für ein rasches Vorwärtswiedergeben sowie auch für die abschnittsweise Audio-Wiedergabe in der Vorwärtsreihenfolge beim Zurückspringen gewählt werden.

30 Bei einem Block 34 wird sodann die hier in Rede stehende Rückwärtsmodus-Wiedergabe, beispielsweise durch Betätigen an einer entsprechenden Taste der Eingabemittel 26, gestartet. Danach wird automatisch bei einem Block 35 abgefragt, ob der

Benutzer eine Gesamtsprunglänge vorgegeben hat, um die im Rückwärtsmodus insgesamt in den Audiodaten Ai zurückgesprungen werden soll. Ist dies nicht der Fall (s. Ausgang N des Blocks 35), so wird automatisch eine entsprechende Rücksprung-Gesamtlänge als sogenannter "Default"-Wert vorgegeben, wie insbesondere ein Zurückspringen bis zum

5 Beginn der Audiodaten, vgl. Block 36 in Figur 8. Hat der Benutzer jedoch eine entsprechende Eingabe durchgeführt (Ausgang Y des Blocks 35), so wird gemäß einem Block 37 die gewählte Rücksprung-Gesamtlänge zugrunde gelegt.

Die so gegebenen zwei Zweige im Ablaufdiagramm von Figur 8 münden sodann wieder zusammen, und bei einem Knoten 38 ist durch einen Pfeil symbolisch

10 veranschaulicht, dass die ab nun folgenden Schritte im Falle einer Textänderung vor der momentanen Audioposition wiederholt werden, ohne dass der Rückwärtsmodus (s. Block 34) beendet wird.

Bei einem Block 39 wird dann abgefragt, ob Spracherkennungs-Ausgangsdaten vorliegen, d.h. ob automatisch mit Hilfe von Spracherkennungssoftware in

15 Spracherkennungsmitteln 17 transkribierte Textdateien Ti vorliegen. Wenn dies zutrifft (s. Ausgang Y von Block 39 in Figur 8), dann wird bei einem Block 40 abgefragt, ob die Option O1 (s. Block 33) vorgegeben wurde. Wenn dies der Fall ist (Ausgang Y von Block 40), erfolgt nachfolgend bei einem Block 41 die Berechnung der Audio-Segmente bzw. Wörter in den Audiodaten Ai in Entsprechung zu den erkannten und editierten Textwörtern

20 in den Textdaten Ti.

Wird jedoch keine Spracherkennung verwendet (Ausgang N von Block 39), so wird bei einem Block 42 abgefragt, ob Textdaten Ti vor der momentanen Audio-Position vorliegen oder nicht. Wenn dies nicht der Fall ist (Ausgang N von Block 42), d.h. wenn noch keine Textdaten Ti vorhanden sind, wird bei einem Block 43 eine Schätzung der

25 Länge von Audio-Segmenten auf Basis der Audiodaten-Gesamtlänge oder des Audio-Energieprofils vorgenommen, vgl. die vorstehende Option O4 bzw. O5.

Liegen jedoch gemäß der Prüfung bei Block 42 bereits Textdaten Ti vor der momentanen Audio-Position vor (s. Ausgang Y von Block 42), so wird daraufhin bei einem Block 44 abgefragt, ob beim Block 33 die Option O2 oder O3 konfiguriert wurde,

30 und wenn eine solche Option O2 oder O3 nicht konfiguriert wurde (Ausgang N von Block 44), erfolgt ebenfalls die vorstehend im Zusammenhang mit Block 43 beschriebene Schätzung. Im Fall des Vorliegens der Konfiguration einer Option O2 oder O3 (Ausgang Y

von Block 44) erfolgt bei einem Block 45 daraufhin eine Schätzung der Länge der Audio-Segmente oder Wörter der Audiodaten Ai auf der Basis der Audiodaten-Gesamtlänge und der Anzahl der Wörter bzw. Silben (Optionen O2 bzw. O3).

In dem mit strichlierten Linien umrandenden Bereich 46 des Ablaufs gemäß der Figur 8 werden nunmehr die eigentlichen Schritte für das Zurückspringen und das abschnittsweise Wiedergeben in der Vorwärtsreihenfolge durchgeführt. Dabei wird bei einem Block 47 die Audiowiedergabe eines Wortes oder Segmentes der Audiodaten Ai in der Vorwärtsreihenfolge und somit in verständlicher Art und Weise gestartet, wobei die Wiedergabe mit der beim Block 33 eingestellten Geschwindigkeit erfolgt. Bei einem Block 48 wird dann abgefragt, ob das Ende des in der Vorwärtsreihenfolge wiederzugebenden Wortes oder Segmentes der Audiodaten Ai erreicht ist. Wenn dies nicht der Fall ist (Ausgang N), wird gemäß einem Block 49 das Wiedergeben fortgesetzt, bis schließlich die Abfrage beim Block 48 ergibt, dass das Ende des Wortes bzw. Segmentes erreicht ist (Ausgang Y von Block 48). Bei einem Block 50 wird sodann in den Audiodaten Ai zur nächsten vorgegebenen Zielposition zurückgesprungen, beispielsweise um eine Rücksprunglänge entsprechend der Länge von drei Wörtern. Danach wird bei einem Block 51 abgefragt, ob die vorgegebene Anfangsposition (vgl. Block 36 bzw. 37) bereits erreicht ist. Wenn dies nicht der Fall ist (Ausgang N), wird zum Block 47 zurückgekehrt. Wenn jedoch die Anfangsposition, d.h. das Ende des Rückwärtsmodus, erreicht ist (Ausgang Y von Block 51), wird bei einem Block 52 der Rückwärtsmodus beendet.

Vorzugsweise ist auch vorgesehen, dass der beschriebene Vorgang, der an sich selbsttätig, durch die Steuermittel 20 gesteuert abläuft, jederzeit vor Erreichen des vorgegebenen Endes manuell durch eine "STOP"-Eingabe an den Eingabemitteln 26 beendet werden kann.

Es sei erwähnt, dass nach dem Korrigieren des Textes unter Umständen eine Neudefinition der Wort-Markierungsdaten Mi oder eine Neuberechnung der Länge der Audio-Segmente nötig sein kann.

Es sei weiters erwähnt, dass nach einem Zurückspringen um mehr als zwei Wörter, etwa vier, fünf, sechs oder mehr Wörter, nachfolgend ein akustisch einwandfreies Wiedergeben von mehr als einem Wort durchgeführt werden kann, beispielsweise von zwei oder drei oder vier Wörtern.

Patentansprüche:

1. Einrichtung zum Wiedergeben von gespeicherten Audiodaten vor, welche Audiodaten zu Textdaten eines aus Wörtern zusammengesetzten Texts korrespondieren, mit Speichermitteln zum Speichern der Audiodaten, in welche Speichermittel zu 5 speichernde Audiodaten in einer Vorwärtsreihenfolge einspeicherbar sind, und mit Speichermitteln zum Steuern des Wiedergebens von gespeicherten Audiodaten in einem Vorwärtsmodus und in einem Rückwärtsmodus, und mit Audio-Wiedergabemitteln, wobei die Steuermittel derart eingerichtet sind, dass sie bei einem Wiedergeben von Audiodaten im Rückwärtsmodus ausgehend von der jeweiligen momentanen Wiedergabeposition in 10 den Audiodaten entgegen der Vorwärtsreihenfolge selbsttätig ein Zurückspringen um eine Rücksprunglänge, die der Länge von mindestens etwa zwei Wörtern entspricht, zu einer Zielposition veranlassen und danach beginnend bei der jeweiligen Zielposition ein Wiedergeben von Audiodaten in der Vorwärtsreihenfolge für nur einen Teil der Rücksprunglänge veranlassen.
- 15 2. Einrichtung nach Anspruch 1, wobei die Steuermittel derart eingerichtet sind, dass sie unter Benützung von den Wörtern zugeordneten Wort-Markierungsdaten als Steuerdaten ein Zurückspringen zur jeweiligen Zielposition veranlassen.
3. Einrichtung nach Anspruch 2, wobei den Steuermitteln Zählmittel zum Zählen von beim Zurückspringen bzw. Wiedergeben erreichten Markierungsdaten 20 zugeordnet sind.
4. Einrichtung nach Anspruch 1, wobei den Steuermitteln ein Zeitsteuerkreis zum Bemessen der Dauer der Audio-Wiedergabe zugeordnet ist.
5. Einrichtung nach Anspruch 1, wobei mit den Steuermitteln Einstellmittel zum Einstellen der Geschwindigkeit der Audio-Wiedergabe verbunden sind.
- 25 6. Einrichtung nach Anspruch 1, wobei die Steuermittel weiters mit Textspeichermitteln zum Speichern von zu den Audiodaten korrespondierenden Textdaten verbunden sind, die mit Textanzeigemitteln verbunden sind, und wobei die Steuermittel dazu eingerichtet sind, dass sie mit Hilfe von Verknüpfungsdaten für die Audiodaten und die Textdaten ein synchrones Wiedergeben der Audiodaten und der hierzu 30 korrespondierenden Textdaten veranlassen.
7. Einrichtung nach Anspruch 6, wobei die Steuermittel und die Textspeichermittel sowie die Speichermittel für die Audiodaten mit einer automatischen

Transkription durchführenden Spracherkennungsmitteln verbunden sind.

8. Verfahren zum Wiedergeben von in Speichermedien gespeicherten Audiodaten vor, welche Audiodaten zu Textdaten eines aus Wörtern zusammengesetzten Text korrespondieren und in welche Speichermedien zu speichernde Audiodaten in einer Vorwärtsreihenfolge eingespeichert werden, bei welchem Verfahren das Wiedergeben von Audiodaten in einem Vorwärtsmodus und in einem Rückwärtsmodus steuerbar ist, wobei bei einem Wiedergeben von Audiodaten im Rückwärtsmodus ausgehend von der jeweiligen momentanen Wiedergabeposition in den Audiodaten selbsttätig um eine Rücksprunglänge, die der Länge von mindestens etwa zwei Wörtern entspricht, entgegen der Vorwärtsreihenfolge zu einer Zielposition zurückgesprungen wird und danach beginnend bei der jeweiligen Zielposition ein Wiedergeben in der Vorwärtsreihenfolge für nur einen Teil der Rücksprunglänge durchgeführt wird.
9. Verfahren nach Anspruch 8, wobei beim Zurückspringen zur Zielposition den Wörtern zugeordnete Wort-Markierungsdaten als Steuerdaten benutzt werden.
10. Verfahren nach Anspruch 9, wobei das Wiedergeben in der Vorwärtsreihenfolge jeweils automatisch beendet wird, wenn beim Wiedergeben die nächsten Wort-Markierungsdaten erreicht werden.
11. Verfahren nach Anspruch 8, wobei das Wiedergeben in der Vorwärtsreihenfolge jeweils automatisch nach einer vorgegebenen Zeit beendet wird.
12. Verfahren nach Anspruch 8, wobei nach Beendigung des Wiedergebens in der Vorwärtsreihenfolge automatisch wieder um eine Rücksprunglänge, die der Länge von mindestens etwa zwei Wörtern entspricht, zurückgesprungen wird.
13. Verfahren nach Anspruch 8, wobei das Zurückspringen in den Audiodaten mit einer gegenüber der Wiedergabegeschwindigkeit beim Wiedergeben in der Vorwärtsreihenfolge höheren Geschwindigkeit und ohne akustisches Wiedergeben der gespeicherten Audiodaten durchgeführt wird.
14. Verfahren nach Anspruch 8, wobei das Wiedergeben der gespeicherten Audiodaten in der Vorwärtsreihenfolge mit einer einstellbaren Wiedergabegeschwindigkeit erfolgt.
15. Verfahren nach Anspruch 8, wobei synchron mit dem Wiedergeben der gespeicherten Audiodaten in der Vorwärtsreihenfolge eine visuelle Anzeige von zu den Audiodaten korrespondierenden Textdaten durchgeführt wird, welches Anzeigen durch

Verknüpfungsdaten für die gespeicherten Audiodaten und die hierzu korrespondierenden Textdaten gesteuert wird.

16. Verfahren nach Anspruch 15, wobei beim visuellen Anzeigen jeweils mehrerer Wörter der Textdaten jeweils jenes visuell angezeigte Wort optisch 5 hervorgehoben wird, von dem die korrespondierenden gespeicherten Audiodaten wiedergegeben werden.

17. Verfahren nach Anspruch 15, wobei die zu Audiodaten korrespondierenden Textdaten durch ein automatisches Spracherkennungsverfahren gewonnen werden, wobei zugleich die Wort-Markierungsdaten als Verknüpfungsdaten für 10 die zueinander korrespondierenden Textdaten und Audiodaten generiert und gespeichert werden.

18. Computerprogrammprodukt, das in einen Speicher eines Computers geladen werden kann und Softwarecodeabschnitte umfasst, um durch deren Abarbeiten nach dem Laden in den Speicher mit dem Computer das Verfahren nach Anspruch 8 15 durchführen zu können.

19. Computerprogrammprodukt nach Anspruch 18, dadurch gekennzeichnet, dass es auf einem computerlesbaren Medium gespeichert ist.

20. Computer mit einer Recheneinheit und einem internen Speicher, welcher Computer zum Abarbeiten des Computerprogrammprodukts nach Anspruch 18 ausgebildet 20 ist.

Zusammenfassung:

Einrichtung und Verfahren zum Wiedergeben von Audiodaten
sowie Computerprogrammprodukt hierfür

5

Beim Wiedergeben von in Speichermitteln (19) gespeicherten Audiodaten (Ai), welche Audiodaten (Ai) zu Textdaten (Ti) eines aus Wörtern zusammengesetzten Texts 10 korrespondieren, wobei das Wiedergeben der Audiodaten (Ai) in einem Vorwärtsmodus und in einem Rückwärtsmodus steuerbar ist, wird für das Wiedergeben der Audiodaten (Ai) im Rückwärtsmodus ausgehend von der jeweiligen momentanen Wiedergabeposition in den Audiodaten selbsttätig entgegen einer Vorwärtsreihenfolge um eine Rücksprunglänge (1A, 2A, 3A...6A), die der Länge von mindestens etwa zwei Wörtern 15 entspricht, zu einer Zielposition zurückgesprungen und danach beginnend bei der jeweiligen Zielposition ein Wiedergeben der Audiodaten in der Vorwärtsreihenfolge für nur einen Teil (1B, 2B, 3B...6B) der Rücksprunglänge durchgeführt.

(Figur 4A)

1/5

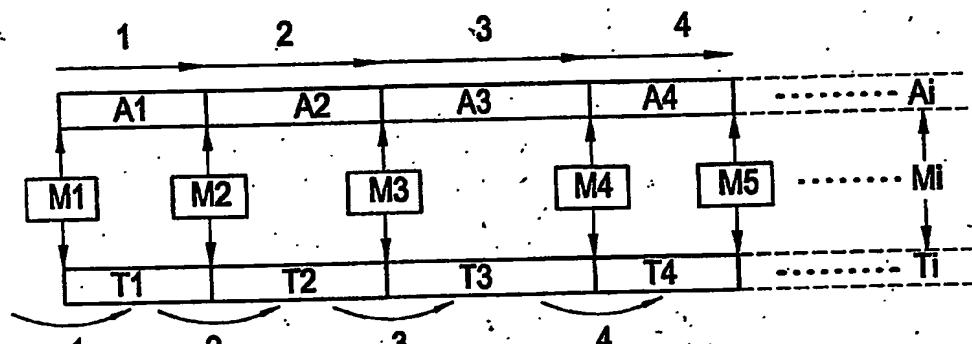


FIG. 1

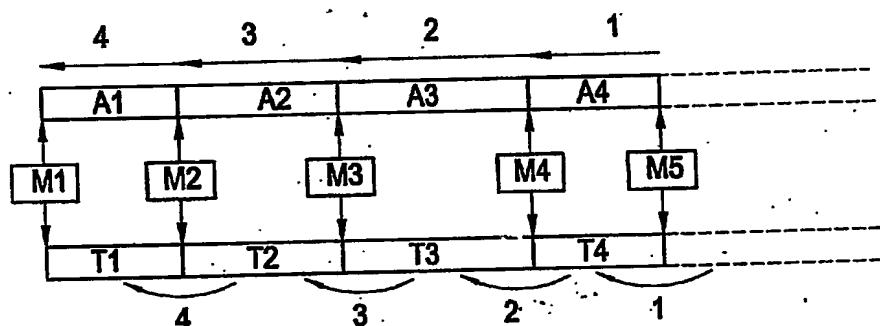


FIG. 2

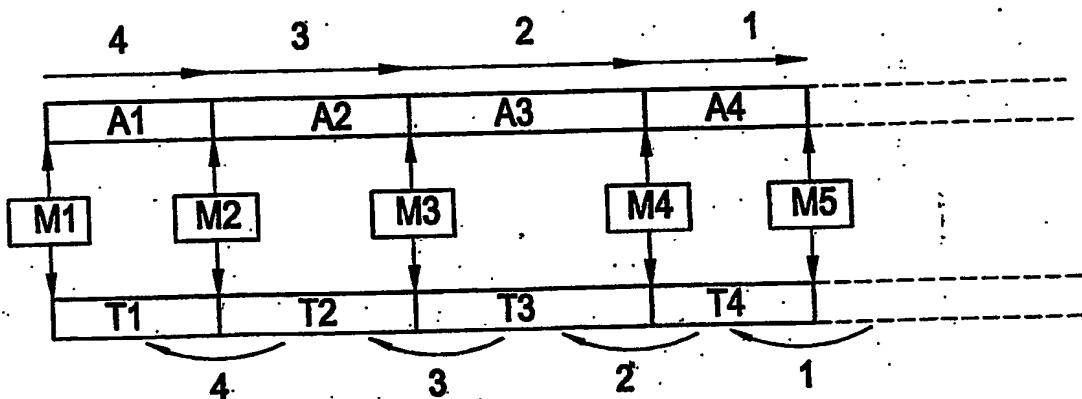


FIG. 3

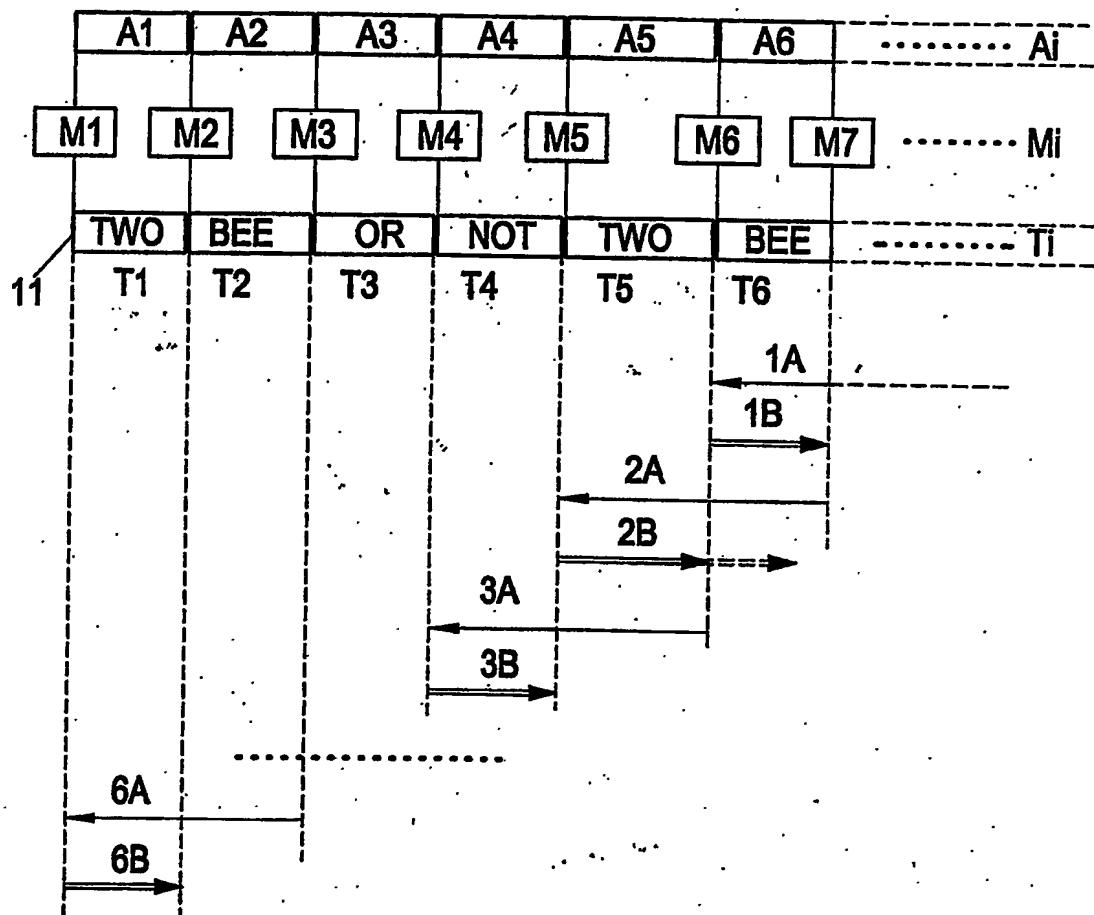


FIG. 4A

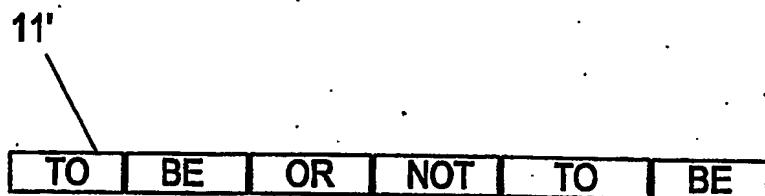


FIG. 4B

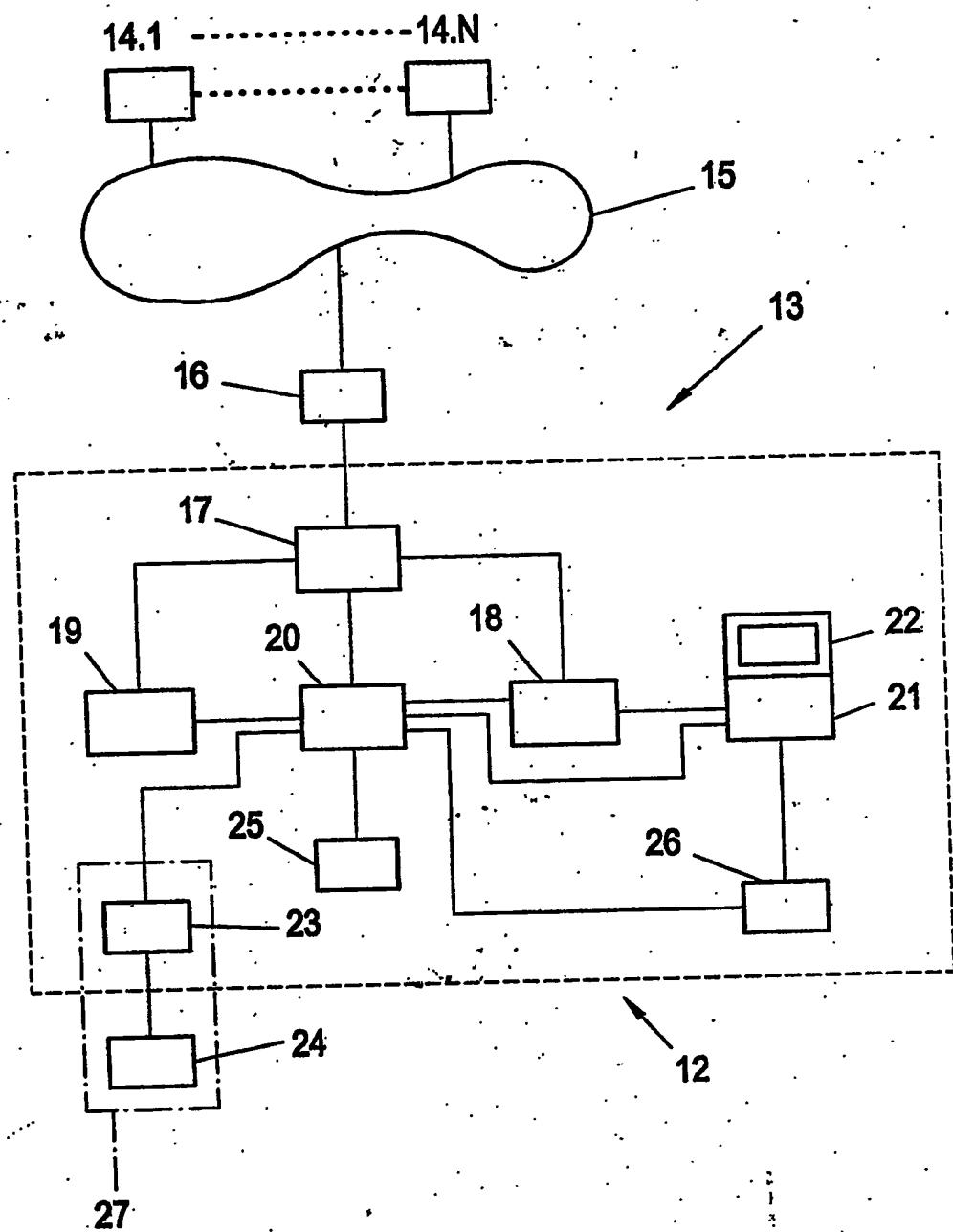


FIG. 5

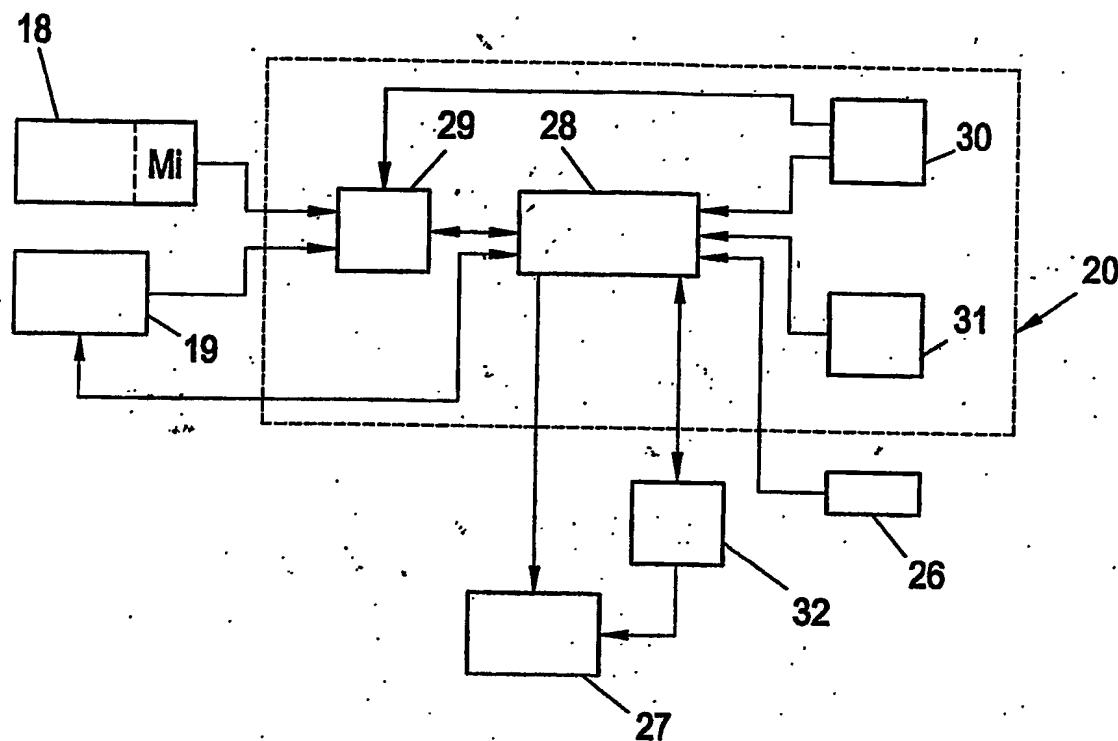


FIG. 6

M1 M2 M3 M4 M5 M6 M7 M8

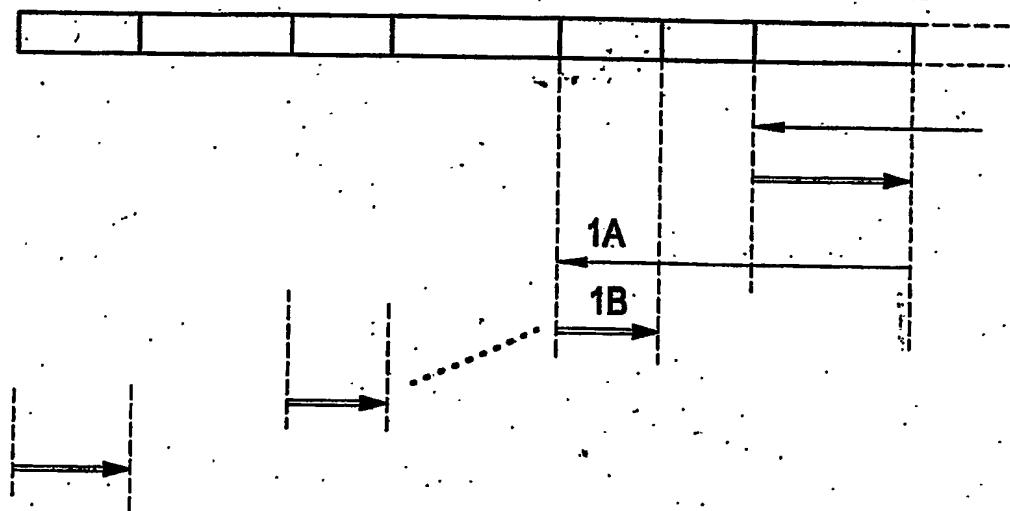


FIG. 7

5/5

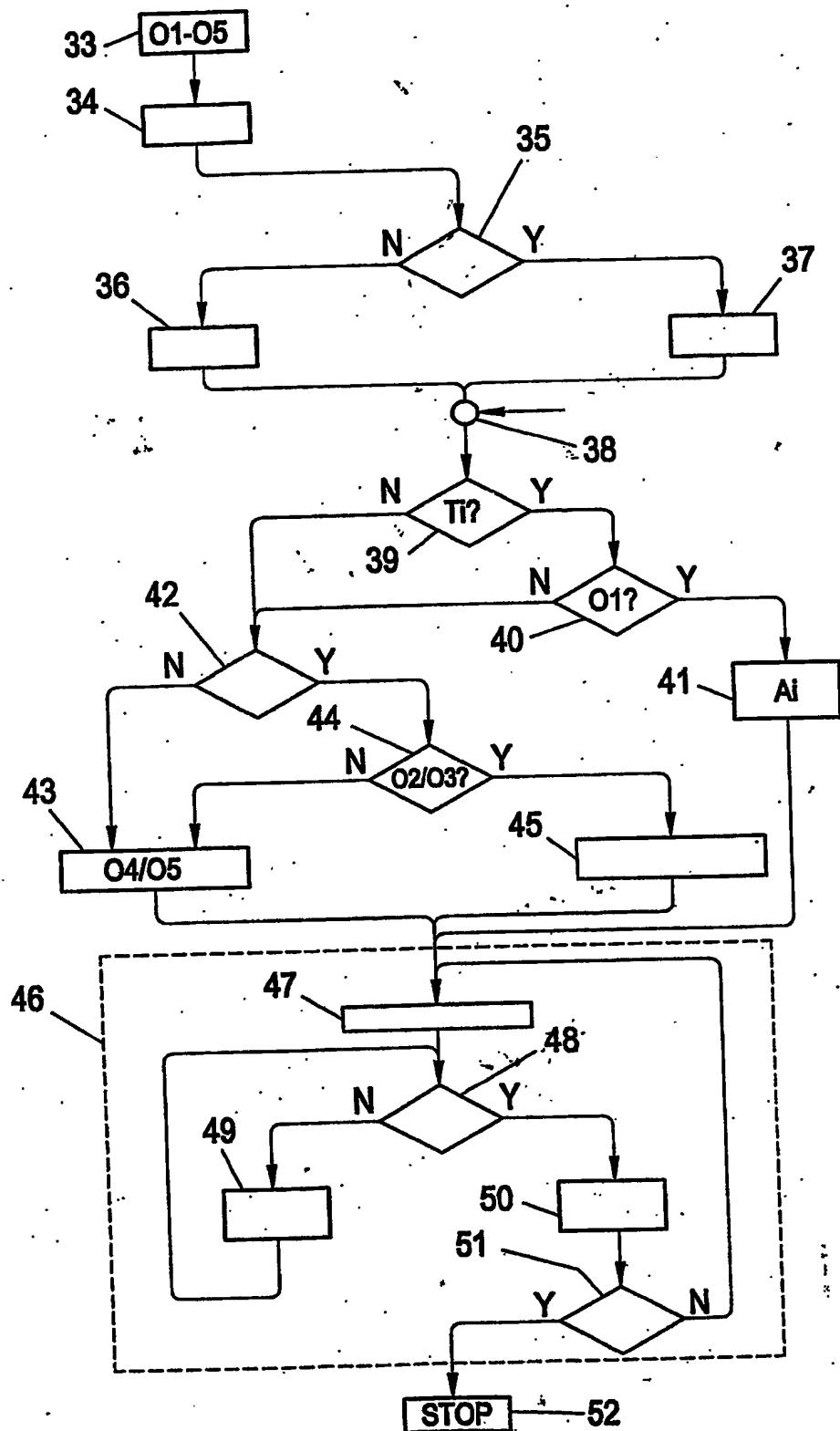


FIG. 8